

AMENDMENTS TO THE CLAIMS

1. (Currently Amended) A voice message processing system, comprising:

a voice message (VM) data store storing voice message data indicative of a plurality of voice messages;

a non-server based voice data processor, configured to be under personal control of a user during runtime voice data processing, coupled to the VM data store, configured to access the voice messages, extract desired information from the voice messages and augment the VM data stored in the VM data store with the desired information, the desired information including speaker identification information, acoustic information, voice message content, and speaking rate information, the augmented VM data comprising: speaker identity of the speaker of the voice message obtained by the voice data processor accessing a plurality of acoustic speaker identification models and obtaining the speaker identification based on the speaker identification information, speaker emotion based on the acoustic information, a voice message summary generated from the voice message content, and a speaking rate obtained from the speaking rate information;

a user interface component coupled to the VM data store and configured to provide user access to the augmented VM data, the user interface component providing a display that displays the speaker identity obtained from the speaker identification models, the speaker emotion, the voice message summary, and an adjustment user interface element that is actuatable to adjust the speaking rate of playback ~~paybaek~~ of the voice message; and

wherein the voice data processor comprises a rate normalization component configured to receive the speaking rate information and normalize an associated voice message to a preselected speaking rate.

2. (Previously Presented) The system of claim 1 wherein the voice data processor comprises:

- a rule application component configured to receive user rule inputs indicative of user-selected rules and to apply the user-selected rules to the augmented VM data.
3. (Previously Presented) The system of claim 2 wherein the voice data processor comprises:
- a speaker identification model data store storing the plurality of speaker identification models; and
 - a speaker identification component configured to access the speaker identification models data store and provide the identity of the speaker associated with the voice message corresponding to the VM data.
4. (Previously Presented) The system of claim 3 wherein the voice data processor comprises:
- a speaker model training component configured to receive VM data and train a speaker identification model based on the VM data and a user input indicative of a speaker of a voice message corresponding to the VM data.
5. (Previously Presented) The system of claim 2 wherein the voice data processor comprises:
- an acoustic feature extractor extracting the acoustic information as acoustic features from the VM data, the acoustic features being indicative of the desired information.
6. (Previously Presented) The system of claim 5 wherein the acoustic feature extractor is configured to extract features indicative of the speaker emotion and provide an emotion output indicative of the speaker's emotion.
7. (Previously Presented) The system of claim 5 wherein the acoustic feature extractor is configured to extract features indicative of the speaking rate and provide a rate output indicative of the speaking rate.
8. (Canceled).

9. (Previously Presented) The system of claim 2 wherein the voice data processor comprises:
a speech-to-text component configured to generate a textual output indicative of a content
of a voice message.
10. (Original) The system of claim 9 wherein the speech-to-text component is configured to
generate a transcription of the voice message as the textual output.
11. (Previously Presented) The system of claim 9 wherein the voice data processor comprises:
a summarization component configured to generate the voice mail summary of the voice
message.
12. (Previously Presented) The system of claim 9 wherein the voice data processor comprises:
a semantic parser configured to generate a semantic parse of at least a portion of the voice
message.
13. (Original) The system of claim 2 wherein the rule application component sorts voice
messages based on the desired information.
14. (Original) The system of claim 2 wherein the rule application component generates alarms
based on the desired information.
15. (Original) The system of claim 2 wherein the user interface component generates a user
interface exposing user-selectable inputs for manipulation of the voice message by the user.
16. (Canceled).

17. (Original) The system of claim 15 wherein the user interface displays a textual indication of a content of a voice message.

18. (Canceled).

19. (Canceled).

20. (Original) The system of claim 15 wherein the user interface displays a rule indicator indicative of rules being applied.

21. (Currently Amended) A method of processing voice messages, comprising:
storing the voice messages at a non-server based voice message (VM) data store;
intermittently accessing the VM data store during runtime to determine whether a new voice message has been stored;
for each new voice message, processing the new voice message at a non-server based processor, during runtime, wherein processing includes extracting acoustic features from the new voice message and accessing acoustic speaker identification models to obtain speaker identity identifying a speaker of the new voice message, the extracted data further comprising acoustic features indicative of other desired information, and a textual representation of a content of the new voice message;
and
augmenting data in the VM data store with the extracted data; and
adjusting a speaking rate of the new voice message to a selected speaking rate for playback.

22. (Original) The method of claim 21 wherein processing the new voice message to obtain acoustic features comprises:

obtaining acoustic features indicative of an emotion of a speaker of the new voice message and generating a speaker emotion output indicative of the speaker's emotion.

23. (Currently Amended) The method of claim 21 wherein the acoustic features include a speaking rate indicator indicative of a speaking rate of the speaker of the new voice message, and further comprising:

providing an interface to enable a user to adjust the normalizing the speaking rate to a user[[-]]selected speaking rate for playback of the new voice message.

24. (Original) The method of claim 21 wherein obtaining speaker identity includes providing an unknown output when speaker identity is determined to be unknown and further comprising:

receiving a user input indicative of a speaker identity for the new voice message; and
training a speaker identification model based on the new voice message and the user input.

25. (Original) The method of claim 21 and further comprising:

receiving a rules input indicative of user-selected rules to be applied to the new voice message; and
applying the user-selected rules based on the extracted data.

26. (Original) The method of claim 21 and further comprising:

semantically parsing the textual representation of the new voice message.

27. (Original) The method of claim 21 and further comprising:

generating a user interface to the VM data store, the user interface including user-actuable inputs for manipulating the voice messages in the VM data store.